

Interactions (Moderators)

STAT 245

Interactions Defined

Two predictors **interact** when you need to know values of *both* in order to make an accurate prediction of the response variable value. One variable *modulates* or *alters* the effect of the other *on the response*.

(Interaction is AKA Moderation!)

Interaction Note

Predictors can interact in *any* type of regression model (so this chapter could really be placed almost anywhere in our course).

Video #1



Video #2



SET Example

- Dataset: `teach_beauty`
- Course evaluation scores (`eval`) for many profs.
- From Gelman -- Columbia?

Dataset includes:

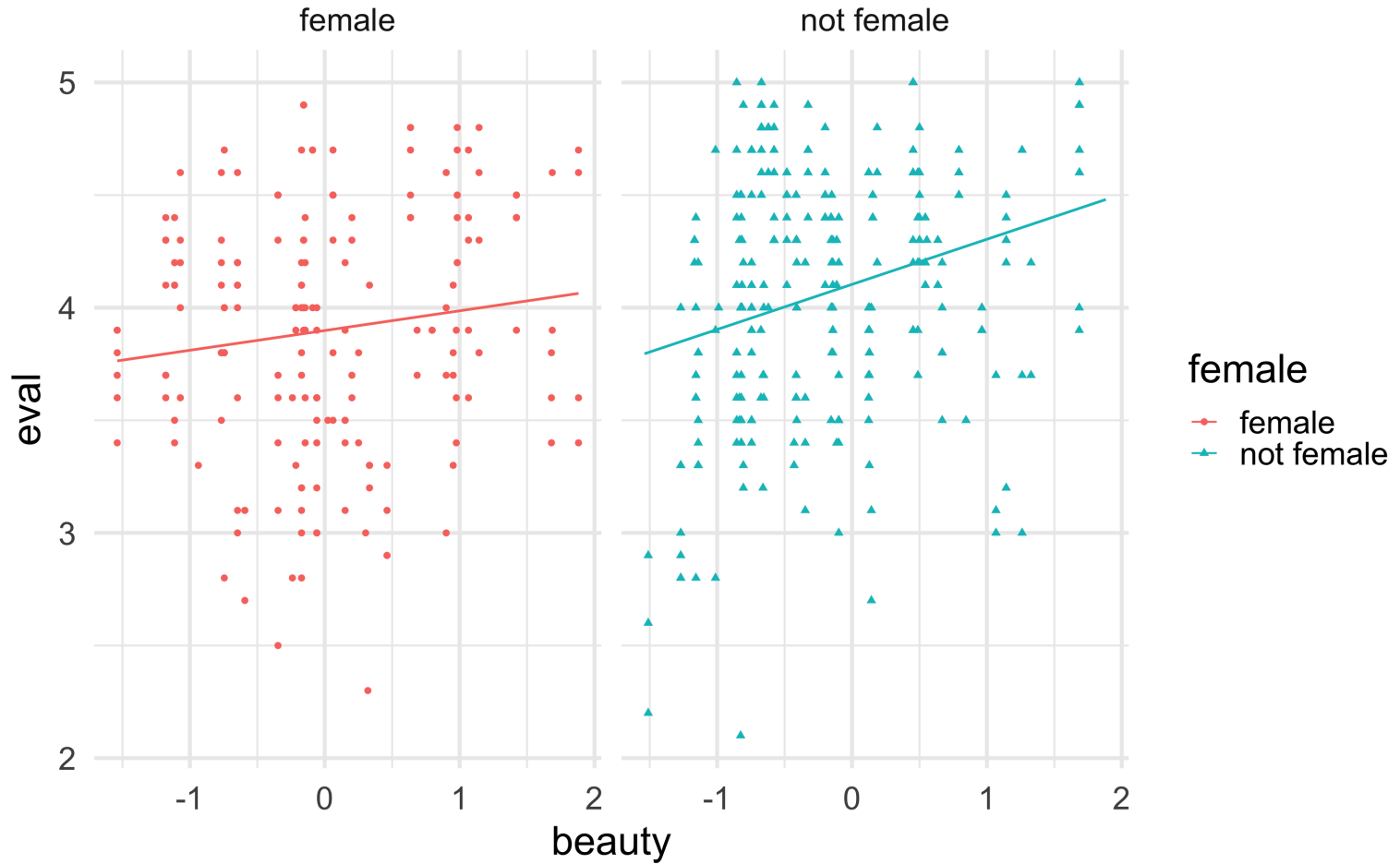
- `native_language`
- `beauty` (?)
- `age`
- is dress `formal`?
- is prof `female`?
- is prof white/caucasian? (`race_eth`)

More Background

- Spoooren et al. 2013,
- Hornstein 2017
- Kreitzer & Sweet-Cushman 2022

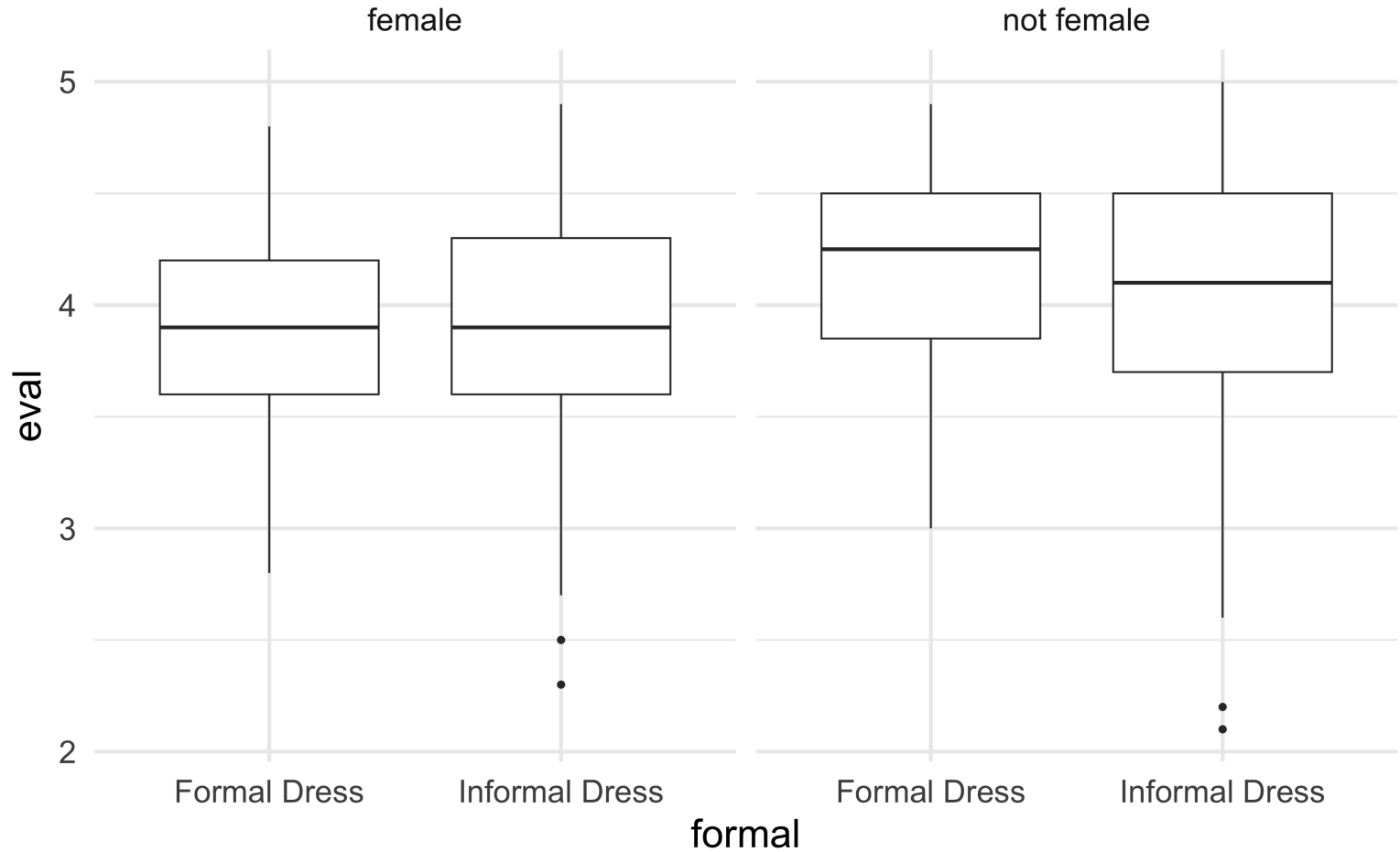
Night Science - We can Peek!

A Categorical-Quantitative Interaction?



Night Science - We can Peek!

A Categorical-Categorical Interaction?



Quant-Quant interactions?

Let's not go there...

Interactions in R Model Formula: * not +

```
beauty_mod <- lm(eval ~  
                  beauty * female +  
                  formal * female,  
                  data = teach_beauty)
```

Consistent Diagram?

summary() w/Interaction

```
msummary(beauty_mod)
```

```
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   3.84839    0.10813   35.591   <2e-16
***
## beauty                        0.09021    0.04744    1.902    0.0578
.
## femalenot female              0.27615    0.13130    2.103    0.0360
*
## formalInformal Dress          0.05751    0.11574    0.497    0.6195
## beauty:femalenot female       0.10841    0.06452    1.680    0.0936
.
## femalenot female:formalInformal Dress -0.08378    0.14276   -0.587    0.5576
##
## Residual standard error: 0.5371 on 457 degrees of freedom
## Multiple R-squared:  0.07326.    Adjusted R-squared:  0.06312
```

summary() w/Interaction -- Equation?

```
msummary( beauty_mod)
```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.84839    0.10813  35.591  <2e-16 ***
## beauty           0.09021    0.04744   1.902   0.0578 .
## femalenot female  0.27615    0.13130   2.103   0.0360 *
## formalInformal Dress  0.05751    0.11574   0.497   0.6195
## beauty:femalenot female  0.10841    0.06452   1.680   0.0936 .
## femalenot female:formalInformal Dress -0.08378    0.14276  -0.587   0.5576
##
## Residual standard error: 0.5371 on 457 degrees of freedom
## Multiple R-squared:  0.07326,    Adjusted R-squared:  0.06312
## F-statistic: 7.226 on 5 and 457 DF,  p-value: 1.587e-06
```

Selection using ANOVA

```
car::Anova(beauty_mod)
```

```
## Anova Table (Type II tests)
##
## Response: eval
##
```

	Sum Sq	Df	F value	Pr(>F)	
## beauty	6.180	1	21.4252	4.8e-06	***
## female	4.303	1	14.9167	0.0001286	***
## formal	0.000	1	0.0013	0.9713449	
## beauty:female	0.815	1	2.8238	0.0935599	.
## female:formal	0.099	1	0.3444	0.5575709	
## Residuals	131.818	457			
## ---					
## Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' ' 1

Selection using AIC

```
m2 <- lm(eval ~ beauty + female + formal * female,  
          data = teach_beauty)  
m3 <- lm(eval ~ beauty * female + formal + female,  
          data = teach_beauty)  
AIC(beauty_mod, m2)
```

```
##           df      AIC  
## beauty_mod  7 746.2676  
## m2          6 747.1197
```

```
AIC(beauty_mod, m3)
```

```
##           df      AIC  
## beauty_mod  7 746.2676  
## m3          6 744.6164
```

Cautionary note

If you include an interaction in a regression model, you **must** also include the corresponding "fixed effects" (It would take effort to mess this up in R).

Another Example: Bees

We will use data from a 2019 paper by Adam Dolezal and colleagues, entitled *Interacting stressors matter: diet quality and virus infection in honeybee health* (<https://doi.org/10.1098/rsos.181803>).

Bees

Honeybee population declines have been linked to multiple stressors, including reduced diet diversity and increased exposure to understudied viruses. Despite interest in these factors, few experimental studies have explored the interaction between diet diversity and viral infection in honeybees... In laboratory experiments, we found that high-quality diets have the potential to reduce mortality in the face of infection with Israeli acute paralysis virus (IAPV).

Data Prep

```
bees <- read_csv('https://sldr.netlify.app/data/bee-virus.csv') |>
  rename(Cage_id = `Cage Number`,
         Virus = `Virus Treatment`,
         Food = `Pollen Treatment`,
         Experiment_id = `Experimental replicate code`,
         Mortality = `72 hpi proportion mortality`) |>
  drop_na(Virus, Food, Mortality)

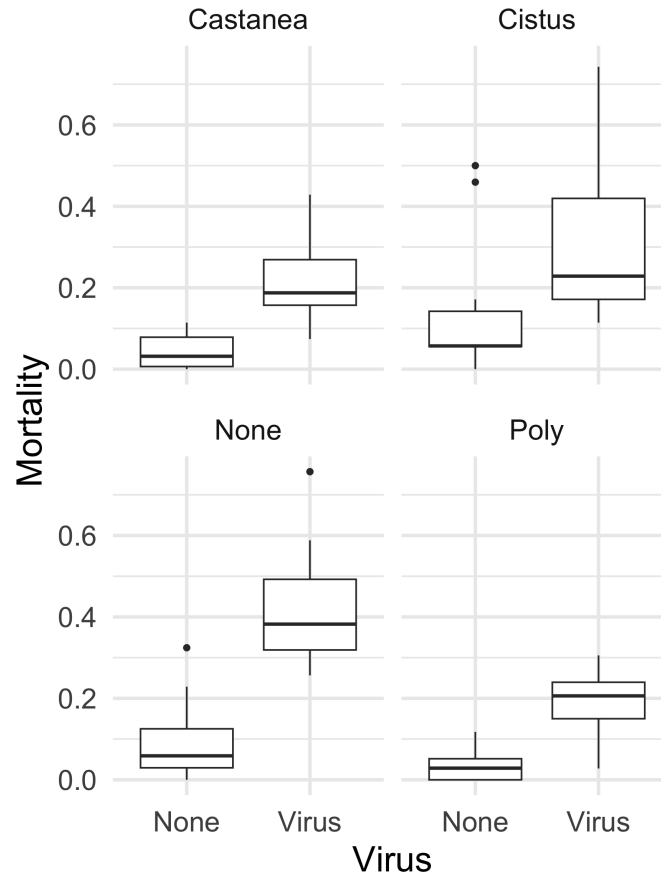
glimpse(bees)
```

```
## Rows: 117
## Columns: 5
## $ Cage_id      <dbl> 351, 352, 353, 354, 355, 357, 359, 360, 361, 362,
363, 3...
## $ Virus        <chr> "None", "Virus", "None", "None", "Virus", "Virus",
"Virus...
## $ Food         <chr> "None", "Poly", "Poly", "Castanea", "Castanea",
"Cistus"...
## $ Experiment_id <dbl> 29, 29, 29, 29, 29, 29, 29, 29, 29, 29, 29, 29,
29, ...
## $ Mortality    <dbl> 0.02857143, 0.11428571, 0.02857143, 0.05555556,
0.171428...
```

Categorical- Quantitative Interaction of and ?

Categorical-Quantitative Interaction

```
gf_boxplot(Mortality ~ Virus | Food,  
            data = bees)
```



Bee Model

```
bee_int_model <- lm(Mortality ~  
                    Virus * Food,  
                    data = bees)  
msummary(bee_int_model)
```

```
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept)    0.04470    0.03202   1.396 0.165533  
## VirusVirus     0.17462    0.04452   3.922 0.000154 ***  
## FoodCistus     0.07966    0.04452   1.789 0.076326 .  
## FoodNone       0.04699    0.04452   1.055 0.293548  
## FoodPoly      -0.01016    0.04528  -0.224 0.822810  
## VirusVirus:FoodCistus 0.01846    0.06241   0.296 0.767981  
## VirusVirus:FoodNone  0.15731    0.06241   2.521 0.013165 *  
## VirusVirus:FoodPoly -0.01704    0.06350  -0.268 0.788894  
##  
## Residual standard error: 0.1198 on 109 degrees of freedom  
## Multiple R-squared:  0.5502,    Adjusted R-squared:  0.5213  
## F-statistic: 19.05 on 7 and 109 DF,  p-value: < 2.2e-16
```

Selection

What's it mean?

```
car::Anova(bee_int_model)
```

```
## Anova Table (Type II tests)
##
## Response: Mortality
##              Sum Sq Df F value    Pr(>F)
## Virus          1.35948  1 94.7322 < 2.2e-16 ***
## Food           0.42364  3  9.8402 8.491e-06 ***
## Virus:Food     0.14103  3  3.2758  0.02385 *
## Residuals     1.56423 109
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Pred Plots

If predictors interact in a model, you shouldn't show a prediction plot with *any* of them unless it shows *all* of them

Pred Plot - How?

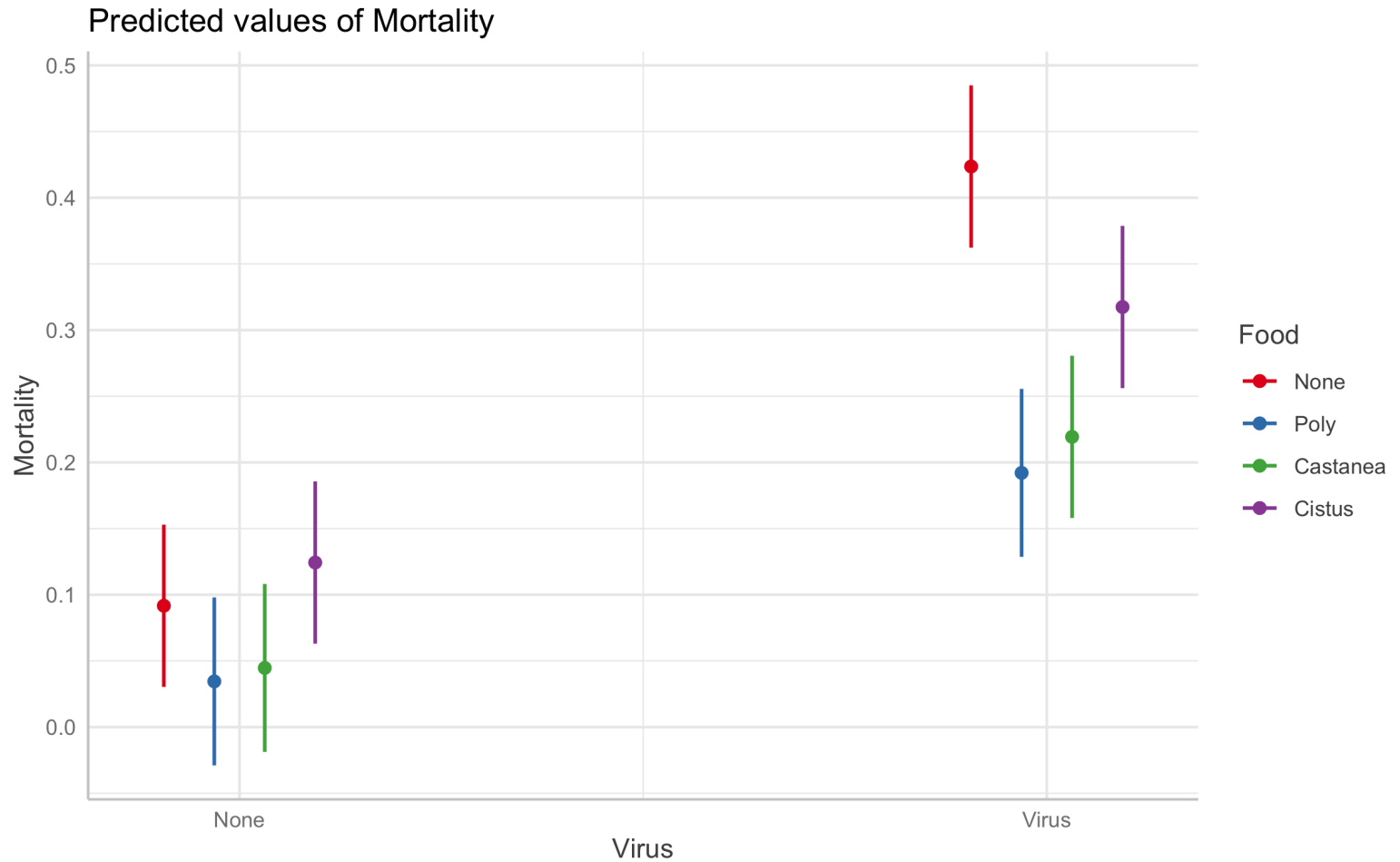
Pred Plot Shortcut

(NOT allowed for use on Test 1!)

```
library(ggeffects)
ggpredict(bee_int_model,
          terms = c('Virus', 'Food')) |>
  plot()
```

Pred Plot Shortcut

(NOT allowed for use on Test 1!)



Pred Plot Shortcut

(NOT allowed for use on Test 1!)

Order matters

```
library(ggeffects)
ggpredict(bee_int_model,
          terms = c('Food', 'Virus')) |>
  plot()
```

Pred Plot Shortcut

(NOT allowed for use on Test 1!)

